Unification and mechanistic detail as drivers of model construction : Models of networks in economics and sociology

Kuorikoski, Jaakko

2014-12

Kuorikoski , J & Marchionni , C 2014 , ' Unification and mechanistic detail as drivers of model construction : Models of networks in economics and sociology ' , Studies in History and Philosophy of Science Part A , vol. 48 , pp. 97-104 . https://doi.org/10.1016/j.shpsa.2014.05.005

http://hdl.handle.net/10138/234271
https://doi.org/10.1016/j.shpsa.2014.05.005

acceptedVersion

UNIFICATION AND MECHANISTIC DETAIL AS DRIVERS OF MODEL CONSTRUCTION: MODELS OF NETWORKS IN ECONOMICS AND SOCIOLOGY

Jaakko Kuorikoski and Caterina Marchionni

Published in *Studies in History and Philosophy of Science* 48: 97-104, 2014.

UNIFICATION AND MECHANISTIC DETAIL AS DRIVERS OF MODEL CONSTRUCTION: MODELS OF NETWORKS IN ECONOMICS AND SOCIOLOGY

Abstract. We examine the diversity of strategies of modelling networks in (micro)economics and (analytical) sociology. Field-specific conceptions of what explaining (with) networks amounts to or systematic preference for certain kinds of explanatory factors are not sufficient to account for differences in modelling methodologies. We argue that network models in both sociology and economics are abstract models of network mechanisms and that differences in their modelling strategies derive to a large extent from field-specific conceptions of the way in which a good model should be a general one. Whereas the economics models aim at unification, the sociological models aim at a set of mechanism schemas that are extrapolatable to the extent that the underlying psychological mechanisms are general. These conceptions of generality induce specific biases in mechanistic explanation and are related to different views of when knowledge from different fields should be seen as relevant.

Keywords. Unification, Mechanisms, Modelling, Networks, Interdisciplinarity

1. INTRODUCTION

Social reality is constituted by the network of relations we have with each other. Properties of these networks influence the way in which social and economic processes start, propagate and influence the agents embedded in them. Network structures are probably the best-understood and least controversial example of the contested concept of social structure. In both economics and sociology network models are increasingly conceptualized as models of social mechanisms. Whereas the sociometric tradition of network research in sociology focuses on the measurement of networks from whatever quantitative data is available, recent sociological network research also investigates how different micro-social mechanisms lead to different network structures and how properties of these structures lead to macro-sociological outcomes. In economics, network theory is a more recent development and it constitutes an attempt to incorporate social structure into economic models, which have been long criticised for their lack of 'embeddedness' (Granovetter, 1983). The economic theory of networks is part of a micro-economic investigation into how strategic individual maximization leads to different network structures and how the study of network properties helps to rectify those empirical failures of standard microeconomics that are due to its ignorance of the social structure of markets. In both fields, the idea of a network as a mechanistic concept is related to the view that macro phenomena are to be explained via the actions and interactions of individual agents. Modelling in the two fields, however, differs significantly in the way in which network theory (the application of graph theory to the study of networked phenomena) is used to build explanatory models (Marchionni, 2013).

This article traces the sources of the difference in modelling strategies in economics and sociology. We show that both fields build network models as models of mechanisms in accordance with the idea of mechanistic explanation, and hence that differences in general conceptions of what explaining (with) networks amounts to do not suffice to account for the diversity of modelling strategies. However, the way in which models of network mechanisms are built, used and evaluated is influenced by the epistemic virtues held dear by each discipline. These conceptions in turn imply different views on when knowledge from different fields should be seen as relevant. The axiomatic ideal of unity in economics implies that modelling assumptions made to account for particular network structures are judged ad hoc unless motivated by general microeconomic theory. This also means that the cognitive mechanisms responsible for the behaviour of individuals are not taken to be a matter of interest for the study of economic networks. In contrast, the idea of unity as relating levels of mechanisms and the idea of mechanism schemas as disjointed parts of a general theoretical toolbox gives more room for manoeuvre with respect to case-specific assumptions – as long as those assumptions have some evidential support, possibly from other sciences.

We use four short case studies to illustrate our claims. Two models of network formation, one from sociology and one from economics, and two models of network structure leading to other interesting macro outcomes, again from a sociological and economic perspective. Our approach is not purely descriptive in that we aim to show that while economic theorizing appears to be in many ways more systematic than the corresponding sociological one, the ideal of axiomatic unity inherent in economics constricts the use of alternative sources of evidence in a way that is in tension with a general causal mechanistic understanding of science.

The examination of this case of network modelling in economics and sociology contributes to the philosophical study of interdisciplinary transfers in two ways: first, it sheds light on how the use of a travelling modelling framework is shaped by discipline-specific epistemic values; and second, it shows that these values affect a discipline's openness towards evidence from other disciplines.

The structure of this article is as follows. Section 2 presents examples of network models in sociology and in economics, which we use to illustrate our claims in the remainder of the paper. In section 3 we argue that all these exemplify models of social mechanisms, or mechanism schemas. Section 4 contains our argument that much of the difference in the way the models are built can be traced to different conceptions of generality and unification. Section 5 discusses the legitimacy of these disciplinary differences and briefly airs our worries concerning biases inherent in such model-based reasoning. Section 6 provides some concluding remarks.

2. MODELLING NETWORKS IN SOCIOLOGY AND ECONOMICS

Models of networks in both economics and sociology either take networks as given and address questions concerning how a person's network of relations influences his/her behaviour, or how different network architectures affect either individual or aggregate outcomes. Alternatively, network models investigate how networks with certain properties emerge as a result of the interactions between individuals.[1] In what follows we first present two models, one from sociology and one from economics, dealing with network creation. We then move on to describe a sociological and an economic model of network influence. Our sample is small and far from random.[2] The models are highly cited in their respective disciplines,[3] and we believe them to be representative of the differences in the methodology we are discussing. We have deliberately chosen these models to illustrate our claims, not to provide empirical inductive support for them.

2.1. Creating networks

The chains of affection model. Network modelling in sociology has become associated with the methodological reform movement of *analytical sociology*, which explicitly aims to provide explanatory (middle-range) theories in terms of social mechanisms. One of the most lauded studies in this genre is the network model of the romantic and sexual relationships in a population of adolescents in a small Midwestern high school (Bearman et al. 2004). The study uses data from the *Add Health* database to map the network of romantic and sexual relationships. It compares the observed structure of the network with a set of simulated networks generated according to alternative models used in the study of disease diffusion. The study finds that the observed network is structured as a spanning tree, exhibiting fewer cycles than would be expected if the network was formed randomly, or simply from individual preferences concerning partner characteristics (such as homophily). Instead, the simulations show that the most parsimonious

[1] In some cases network formation and network influence are investigated in the same model.

[2] See also Marchionni (2013), which discusses Jackson & Wolinsky (1996) and Bearman et al. (2004), and Marchionni & Ylikoski (2013), which takes Centola et al. (2005) as a case study.

[3] To give a sense of their popularity we checked the number of citations on Google Scholar and on Scopus (Elsevier): Jackson & Wolinsky (1996): 1541 citations in Google scholar and 443 in Scopus; Bearman et al. (2004): 344 in Google Scholar and 143 in Scopus; Ballester et al. (2006): 364 in Google Scholar and 70 in Scopus; Centola et al. (2005): 91 in Google Scholar and 34 in Scopus. Both databases have been accessed on 12. March 2013.

micro-mechanism from which such a network structure can be generated is a simple rule forbidding cycles of four (‘don’t date your old partner’s current partner’s old partner’), combined with other reasonable preferences governing partner choice (homophily). The simulation model provides a possible micro-level behavioural rule that would result in a spanning tree structure. It is not assumed that the behavioural rule would be a norm explicitly held in the population such that the high school students would be able to recognize and articulate it. The researchers reason that a relationship completing a cycle of four would constitute a ‘seconds partnership’ and would thus lead to a loss of social status. The modelling methodology thus starts from empirical data from a single population and uses computer simulation to infer to an individual level behavioural rule that could have produced the observed network structure. This behavioural rule is then linked to empirical hypotheses concerning the socio-psychological mechanisms leading to such behaviour. The potential for the extrapolation of the network formation mechanism lies in the generality of the socio-psychological mechanisms.

The connections model. In contrast to sociology, networks have only recently become a distinct object of analysis in economics, with an identifiable body of substantial theoretical and empirical literature (Goyal, 2007; Jackson, 2008). Jackson and Wolinsky (1996), a foundational paper of economic network formation, models the network as the result of individual strategic decision-making and studies the conditions under which alternative network architectures are stable. The model assumes that each agent faces the choice of whether to connect with other agents. Creating a connection requires a certain amount of effort (cost), but each connection is valuable in that it allows access to whichever benefits are available to the agent to whom one is connected. Although only direct connections are costly, both direct and indirect connections are valuable for the agent. Hence, the agents decide to form a connection when the benefits are larger than the costs. As usual for economic models, the networks that are eventually created are those that are stable in terms of equilibrium. In this model, equilibrium is defined in terms of pairwise stability: a network is pairwise stable when no agent has the incentive to undo any of the existing connections, and no pair of agents has the incentive to form a non-existing connection. Depending on the particular configuration of costs and benefits, different network architectures are then shown to be stable and/or efficient. Here the modelling effort does not begin from an observed network as in the sociological model presented above, but from the abstract description of an optimization problem.[4] Unlike the *chains of affection model*, the connections model is solved by analytically deriving an equilibrium solution for the strategic optimization problem. Partly this reflects economists' preference for analytical models compared to simulation models (see Lehtinen & Kuorikoski, 2007). In addition to the *connections model*, the general framework is then applied to a co-authorship situation (i.e. the co-author model). However, the function of the co-author model is to demonstrate the analytical power of the general modelling framework, not to provide (anecdotal) empirical evidence for the modelling assumptions.

2.2. Explaining with networks

The emperor's dilemma model. Not all norms are functional. Sometimes a conflict exists between private attitudes and public behaviour: the majority is willing to stop a practice, but falsely believe that the majority is still willing to accept and enforce it and therefore continue endorsing the unpopular norm. But how can such privately unpopular norms become fixated in a population? Centola, Willer and Macy (2005) investigate the emergence of such unpopular norms with a network simulation. The model is built from a population of heterogeneous agents who have to decide whether or not to comply with and enforce a norm. The norm is supported by a few *true believers* but opposed by the majority. True believers always comply with the norm and when dissatisfied with the level of compliance by others, they also enforce the norm. Disbelievers can be pressured to comply with the norm and even to enforce it. At every iteration of the simulation, agents observe how many neighbours (nodes they are directly linked to) follow the norm and how many enforce it (level of public support for and against the norm). Agents then decide whether to comply and whether to enforce. The central explanatory hypothesis is that false enforcement is used as a signal of (apparent) sincerity and that such signalling can result in cascades leading to the fixation of the unpopular norm. The more the network allows agents to correctly track the level of support of the norm, the less likely it is that a cascade will be triggered. The structure of the

[4] This applies to most economic models of network formation, but there are also exceptions. See e.g. Currarini et al. (2009).

network is a crucial factor in the formation of cascades. Cascades fail to be triggered in a fully connected population, in an embedded population with a small number of randomly distributed true believers and when random ties reduce the overlap between local neighbourhoods. The model relies on two individual behavioural assumptions. The first is that conformity to an unpopular norm can lead to false enforcement in the absence of an explicit metanorm. The second is that false enforcement is an effective signal of apparent sincerity. The accuracy of both behavioural assumptions is tested in an experimental follow-up study (Willer et al. 2009). Additionally, the model is motivated by the psychological finding of an 'illusion of transparency': people systematically overestimate the extent and accuracy of other people's views of their private beliefs and desires (Gilovich et al., 1998; Savitsky et al., 2001). This cognitive bias is theorized to be a factor in the perceived need for reliable signals of sincerity.

The key players model. The productivity of many of our activities is dependent on the activities of those around us. Such peer effects have traditionally been modelled simply as frequency-dependence. In order to capture the consequences of network structures for peer effects, Ballester et al. (2006) built a network game with local payoff complementaries. The model assumes an n-player game with linear quadratic payoffs, with each player simultaneously selecting an effort (e.g. in criminal activity or education). Agents are connected through a network of peer influences and they reap complementarities from their direct network peers. Ballester et al. showed that if the direct payoff from the agent's own effort is concave enough so as to counter the network complementaries, there is a unique Nash equilibrium in which each agent's strategy is proportional to his/her Katz-Bonacich centrality.[5] Bonacich centrality measures the number and weight of both direct and indirect connections of an individual, where less weight is given to more indirect connections. In equilibrium, each player contributes to the aggregate production in proportion to this centrality measure (e.g. the more central/connected an individual is in the network, the higher his level of criminal activity) and the aggregate equilibrium outcome increases with the density and size of local interactions (e.g. the denser and bigger the network, the higher the aggregated level of criminal activity). Another interesting result of the model is that the most central player might not be

[5] This measure of centrality was developed in social network analysis.

the most crucial player in terms of affecting the aggregate production of the network. The model allows the identification of key players, based on a different measure of centrality: intercentrality. In addition to the Katz-Bonacich centrality, intercentrality also includes a node's contribution to the Bonacich centrality of all the other nodes and thus captures the externalities created by peer effects. When a parameter describing the decay factor (the extent to which the weight of indirect connections diminishes) is high, the central player and the key player do not coincide, and the removal of key players is the best strategy for influencing the aggregate network production. For example, removing the criminal with the most connections may not be the most effective way of reducing the operations of a crime ring. Instead, the best strategy is to remove the node that contributes most to the connectivity of other nodes. In a follow-up study Calvó-Armengol et al. (2009) tested the postulated relationship between the Bonacich centrality and the individual level of effort in education using Add Health data and found that an increase in the Katz-Bonacich centrality did increase the level of school performance. However, this is a test of the implications of the model and not an independent test for the plausibility of the behavioural assumptions. Although peer effects is a much investigated psychological phenomenon, it is here defined in such a general sense as to render such psychological research largely irrelevant for the network mechanism.

3. MODELS OF NETWORKS AIM AT MECHANISTIC EXPLANATION

It might be the case that the models presented above are different because they were designed for different epistemic ends. After all, a venerable tradition in economic methodology claims that economic models are designed purely for prediction, whereas such an instrumentalist stance is quite alien to most sociologists. It has also been argued that mainstream economics is committed to the deductive-nomological model of explanation (Kincaid, 2012; Lawson, 2003), whereas covering law explanations are certainly not taken as the gold standard in sociology. In contrast, as we will show, the models presented above are all models of social mechanisms built for the provision of mechanistic explanations.[6] If we are right, then the *differences in the modelling methodology do not derive from radical differences in epistemic and explanatory aims*, such as

[6] We do not claim that all network models in either discipline are models of mechanism. Much of network modelling in sociology, especially in the past, has aimed at purely measuring and describing the properties of different patterns of relations in the populations studied.

aiming at prediction instead of explanation, or aiming at causal explanation instead of thick description and interpretive understanding. Much of the recent methodological discussion in the social sciences has emphasized the importance of mechanistic explanation and called for its widespread adoption as a panacea for perceived methodological ills. Still, adherence to the idea of mechanistic explanation leaves room for significant differences in the style of modelling and theorizing.

3.1. Mechanism schemas in the social sciences

We take *mechanisms* to be structures performing a function in virtue of its component parts, component operations, and their organization (Bechtel & Abrahamsen 2005, p. 423). A *mechanistic model* shows how the orchestrated functioning of the mechanism is responsible for one or more phenomena.[7] These definitions are enough for our purposes in that we are not after the one correct analysis (or metaphysics) of the notion of mechanism. What matters to us are the distinctive (epistemic) properties and functions of mechanistic models and mechanism-based explanations.

First, mechanisms are always mechanisms *for* something, i.e., they are identified by the causal role they play in some larger context. Second, mechanism is a causal notion and the relevant properties of the parts are causal properties. Third, mechanisms are structured and this structure is not transparent in the data that a mechanistic model is used to account for (data from which the existence of the explained phenomenon is inferred). Thus mechanistic explanations 'open the black box' underlying some observed regularity. But note that this does not make mechanisms unobservable by definition, just that the causal assumptions behind the mechanistic hypotheses require a different ('lower level') kind of evidence than the regularity to be explained.

This relates to the fourth important point: mechanisms form hierarchies in that the relevant causal properties of the component parts are in turn realized by mechanisms. An implication of this latter point is that mechanistic concepts guide the way in which knowledge from different disciplines

[7] In order for a model to qualify as mechanistic, it need not include an explicit formalization of the dynamics of the interaction of the parts. It is enough that the proposed interpretation of the formal framework includes an account of how the causal interaction of the parts brings about the model result. We take up this point again later in the paper.

should be related. Knowledge from neighbouring fields is relevant insofar it constrains the space of possible mechanisms responsible for some phenomenon of interest: ‘lower level’ evidence is relevant to the extent to which it informs the assumptions concerning the relevant causal properties of the parts and ‘higher level’ evidence is relevant when it provides ecological and organizational constraints and new hypotheses about the possible causal roles of the system of interest. Such *levels of mechanism* are contextual in that what kinds of things are components and what properties of these things are relevant is always relative to the way in which the mechanism is defined. Therefore levels of mechanisms may not neatly correspond to domains of phenomena of ‘disciplines’ or to fields of science (Hedström & Ylikoski, 2010). The locality of levels of mechanism means that the evidential and explanatory relevance of findings from neighbouring disciplines cannot be taken for granted, or summarily dismissed, on a general level. Such piecewise, inter-field relations of relevance among levels of mechanism constitute a particular picture of unity of science, which Carl Craver has dubbed ‘mosaic unity’ (Craver, 2007, pp. 190-195, 228-271). Paradigmatic mechanisms are intentionally designed contraptions with easily distinguishable and functionally modular parts performing different subtasks required for the production of the behaviour of the whole. In such cases, opening the black box is simply a matter of pairing these component operations with different component parts. Most of the philosophical literature on mechanistic explanation and research heuristics in the life sciences discuss ways of investigating such componential mechanisms (e.g., cell mechanisms).

In the social sciences, the hard work in building mechanistic models is not that of finding out the proper component parts, as these are usually agents, but that of figuring out what properties of the parts are important and how the way in which the parts and their properties are organized and interact produces the behaviour of the whole. It follows that in the social sciences the identity of a given mechanism is not so dependent on the exact causal nature of the parts, i.e., what mechanisms realize these causal properties, but on the specific form of organization and interaction that is relevant to the modelled macro phenomenon. Schemas or kinds of social mechanisms are better conceived as *abstract forms of interaction* (Kuorikoski, 2009). In this conception, a mechanism type can be realized by any number of different causal bases. Thus, as it will become clearer below, mechanistic models in the social sciences look different from much of the life sciences. However, this does not make them less mechanistic insofar as they fulfil the same epistemic and explanatory roles we have discussed above.

In line with Craver's idea of mosaic unity, Hedström and Ylikoski (2010) suggest that general mechanism schemas provide a conception of unity of the social science that also contrasts with the ideal of a grand unified theory. Mechanism schemas constitute a general, but quite possibly disjoint and fragmented toolbox. The same mechanism schemas can be used in different social scientific disciplines and different fields can contribute to the toolbox by informing and constraining the mechanism schemas in accordance with the idea of levels of mechanism mentioned above (Hedström & Ylikoski, 2010, p. 61).

3.2. Network models as mechanism schemas

Both sociological models are models of social mechanisms in the sense proposed above. They describe how the behaviour of parts and their interactions produce a certain phenomenon. The *chains of affection* model describes how two assumptions about individual behaviour, homophily and the no-four-step-cycles-rule, produce a spanning-tree network rather than a core structure (the phenomenon to be accounted for). These behavioural assumptions are causal: they describe what the components of the investigated social system do. In turn, a suggestion is made that the behaviour of the components results from socio-psychological mechanisms related to partner choice and status-loss aversion. The *emperor's dilemma* model shows how the properties of the network structure combined with different behavioural rules of the embedded agents are related to the probability of cascades in the network. The behavioural profiles (disbelievers, opportunists and true believers) and their social relations are the causal micro properties, which constitute the network-level disposition for cascading. Assumptions about these causal properties are supported by prior and model-inspired psychological experiments. These models explain network-level properties in terms of the causal properties of the nodes together with their organization.

Although the style of model building differs in several other respects, the two economic models also aim to explain system-level phenomena in terms of the properties of the parts and their organization. The *connections* model provides a template or a mechanism schema in which properties of the parts (preferences, beliefs and costs of forming links) can be plugged into and which relates such attributions to different stability properties of the ensuing networks. The model aims to capture a mechanistic explanatory dependency between the agents' actions and the macro-level outcome: the idea is that the network would not be created if the agents did not base their decisions about forming links on such considerations.[8] In turn, the behaviour of the agents is dependent on changes in incentives. The *key-players* model locates the components that are most crucial for the total production of a network. It describes how the structure of the network, and in particular an agent's position in it, affects his or her decision concerning the level of effort, and how this in turn is responsible for the total level of activity. The model explains system-level properties in terms of the causal properties of the nodes (again, their preferences, beliefs and costs of forming links) together with their organization. Both are analytically solved equilibrium models that do not explicitly model the dynamics leading to equilibrium. Nevertheless, such equilibrium models can be mechanistic in that their interpretation relies on the accompanying narrative (see, e.g., Morgan, 2001; Grüne-Yanoff & Schweinzer, 2008), which describes how the interaction between the agents brings about the equilibrium state (Kuorikoski, 2007). It is only through the narrative that the formal apparatus becomes a model of something, which can be used to yield explanatory insights. Finally, even more clearly than in the case of the sociological models, in these economic models the identity of the mechanism is fixed by the *abstract form of interaction* in that the assumptions about the agents' behaviour are not chosen to fit the situation the model is meant to depict, but derive from the economics framework.

3.3. Filling in the mechanism schemas

If sociological and economic models of networks (at least those roughly of the kind presented above) are models of mechanisms and take the behaviour of individuals—based at least partly on the beliefs and desires attributed to the agents and their interaction through their network of social connections—as the building blocks, why are there dramatic differences in modelling methodology?

[8] What matters is that agents behave in accordance with incentives. How that is actually realized, that is, whether agents go through the rational calculations or not, does not affect the explanatory import of the model. We thank an anonymous referee for pressing us to clarify this point.

A possible explanation is that economists and sociologists *systematically pick out different properties of the components of the network mechanisms (viz. the agents) as relevant*. Something like this has been and still remains a popular way of conceiving the division of labour between sociology and economics (e.g. Coleman, 1988): sociology studies the macro consequences of norm-based behaviour while economics is the science of maximization of self-regarding (often material or monetary) utility. Economists de Martí and Zenou (2011) have recently characterized the approaches to networks in sociology and economics along similar lines:

> Most sociologists would explain (most) networks as an *unintended* outcome of other kinds of activities that individuals engage in. Individuals grow up in certain neighborhoods, they attend certain schools, they take jobs at certain workplaces, etc., and as a by-product of this, they get friends and acquaintances that become nodes in their networks. Their choice of friends and acquaintances rarely is based on their instrumental usefulness. On the contrary, economists would explain networks as an *intended* outcome stemming from strategic interactions (de Martí & Zenou, 2011, p. 340; emphasis in the original).

One may then expect that explaining networks as either unintended or intended outcomes requires different kinds of models. This is not so, however. Although it is true that both sociological models reviewed above are models of norm-governed behaviour, in neither is the norm strictly an exogenous variable that controls the behaviour of the agents, as de Martí and Zenou would seem to imply. Instead, in both models the norms arise endogenously from the choices of individual agents. Moreover, in the *connections* model, the utility gained and the costs incurred in the formation of links can be almost anything, including things such as social status or the displeasure of witnessing unwholesome behaviour. The *key-players* model is explicitly conceived of as a contribution to the peer-effect literature in sociology. So, there are no reasons in principle why one could not model partner choices and norm enforcement in an economic modelling framework. Similarly, one could model co-authorship relations or criminal networks with a network simulation based on case-specific social-psychological assumptions. Therefore, although the properties of agents typically picked out in economics and sociology are not of the same kind, the differences in modelling methodologies cannot be fully accounted for by their selection of kinds of properties.

4. UNIFICATION VS. MECHANISTIC DETAIL AS DRIVERS OF MODEL CONSTRUCTION

All sciences aim at some form of generality. Systematic knowledge cannot be about absolute particularity. But the specific way in which generality is conceived and what is taken to be the privileged form of generality varies across disciplines. By generality we mean the set of actual or possible systems that the model 'applies to' (Matthewson & Weisberg, 2009). We do not delve into what exactly is required for applicability. It suffices to say that more than simple fit with data is required (the model has to correctly 'say' something about the system). A model or theory is more general the more actual or possible systems it applies to. On top of this, unification also requires that the model or theory should be as parsimonious as possible. A model or a theory is unifying if it explains many things, which were previously conceived of as being of a different kind, as manifestations of the same parsimonious set of principles.

There are two ways in which the two sociological network models can be seen as being general. First, abstract network mechanisms are general simply by being abstract: whichever entities satisfy the relevant behavioural assumptions (no-four-step-cycles rule and homophily; enforcement as a signal of sincerity, etc.) will also instantiate the abstract network mechanisms. The mechanism schema applies to many social systems because it is undemanding of what kinds of things the system has to be composed of in order to instantiate it. Second, the models are also general in that the relevant assumptions are based on psychological mechanisms, which are in turn expected to be general. For example, the *chains of affection model* does not make any psychological assumptions that have to be specific to the student population in the particular Midwestern high school. The scope of the network mechanism is determined by other conditions, such as the possibility for actors to watch each other, their capability of recording previous partnerships and their susceptibility to collective assessment of their choices.

In fact, neither sociological model is built according to any *general theory of action*. The chain of affection model infers a behavioural rule from specific network data and comparative simulations. The *emperor's dilemma model* is built on simple psychological assumptions that are not implemented in the simulation model according to any standardized general formalization. From the economist's point of view, such theorizing would be considered ad hoc – what have we gained if we demonstrate that a puzzling macro-behaviour can be generated from equally puzzling micro properties? Such models can also seem like just-so stories in that the behavioural assumptions are simply selected so as to accommodate the model to the dataset or observed phenomenon to be explained. The psychological assumption of avoiding 'seconds partnerships' and the associated loss of social status in the *chains of affection model* is proposed specifically as an explanation of the no-cycles rule and is not grounded in independent psychological research. Therefore, it comes with a certain flavour of adhocness. As Bearman et al. themselves notice, jealousy or the 'yuck factor' (avoidance of too much closeness) are other potential micro-mechanisms (Bearman et al., 2004, p. 75, ff. 35). That the models are not built upon any standard modelling framework also makes comparing them to similar models more difficult. This in turn may appear to hinder the cumulative growth of model-based knowledge of network mechanisms.

Yet, consider the *emperor's dilemma model* again. Therein the causal assumptions are far from arbitrary but are grounded in psychological research. Moreover, a series of follow-up psychological experiments have been performed in order to test the key assumptions of the model. One study aimed at testing the psychological mechanism that leads from conformity to false enforcement, participants tasted and evaluated three wines and also evaluated the performance of other tasters. All the samples were from the same bottle, but one was tainted with vinegar. The experiment included one simulated 'deviant', all ratings were public, but judgements of the performance of others were public only for half of the participants. Subjects favoured the conformist targets over the deviant only when they themselves had conformed and only when peer judgements were public. In private, all favoured the deviant. A second study tested the robustness of false enforcement under different conditions. In this study, subjects were asked to read an unintelligible academic text, ostensibly written by a respected academic. The third study was designed to test the hypothesis that norm enforcement is a credible signal of sincerity. Here the subjects were told that a person had stated publicly that the unintelligible text was clear. Some were told that the person had given public negative evaluations of dissenters whereas others were not. Those who were told of the enforcement were more likely to accept the evaluation as sincere (Willer et al. 2009).

The creators of the *emperor's dilemma model* therefore take the idea that mechanistic modelling should relate evidence from different 'levels' very seriously. Levels of mechanism are local, and the causal assumptions about the properties of the component parts also have to be, in principle, established locally. This is to say that findings or theory from a 'lower' level (in this case psychology) cannot be imported (or ignored) without considering the exact way in which the particular (social) mechanism, and consequently its component parts and their relevant properties, are defined. In the case of the *emperor's dilemma model*, this was done by carrying out additional psychological experiments and by appealing to existing experimental results. The idea of hierarchy of levels of mechanism informs what kind of knowledge from neighbouring disciplines is and should be relevant for sociological questions.

By contrast both economic models use game theoretical solution concepts in order to derive very general properties for network formation and network structure. Here model building does not start from any particular, well specified, network phenomenon and the assumptions concerning the relevant properties of the nodes are not argued for in any way. Instead, the properties attributed to the nodes are dictated by the general axioms of rationality. This makes the models general in two ways. First, like the sociological models, the abstract network mechanisms are general because they are abstract: whatever things satisfy the behavioural constraints formalized in the solution concepts will instantiate the network mechanisms. Second, unlike the sociological models, the economic models are part of a unifying framework, that of economic theory, which strives to account for as many social phenomena as possible with as parsimonious a set of axioms as possible. From this perspective, the use of mechanism-specific assumptions about the causal properties of the parts/agents is not motivated unless they can be derived from this privileged set of axioms. As long as the assumptions are in line with the unifying theory, there is no need to inquire after lower-level (psychological) mechanistic explanations of the properties of the parts. The axiomatic, unifying structure of theorizing makes model comparison straightforward, at least in principle, since all between-model relations are theorems. One can simply check mathematically what other assumptions different models make in addition to the basic axioms and whether some models are special or limiting cases of others. This creates the appearance of a cumulative growth in model-based theorizing.

The ideal of unification is in principle compatible with the idea of mechanistic explanation in that abstract mechanism-schemas can be defined so that the schema selectively picks a certain configuration of causal properties (properties that happen to obey the unifying axioms) in as many situations as possible. Nevertheless, there is a tension between favouring unifying mechanism schemas and the idea of the integration of knowledge according to levels of mechanisms. If a large premium is given for a mechanistic model because of its compliance to a set of unifying principles, then investigation into properties and mechanisms falling outside this privileged set is discouraged. In its extreme form, this leads to the idea of a separate science, which, by definition, is only concerned with those aspects of social systems that are compatible with its analytic machinery. Not surprisingly, some philosophers have suggested economics to be precisely such an enterprise (Hausman, 1992; Ross, 2005). The idea of a separate science is in sharp contrast to mechanism-based ideas of the unity of science, such as that mechanism schemas could form a common but disjointed toolbox for the social sciences (Hedström & Ylikoski, 2010) or the conception of mosaic unity based on the idea of levels of mechanism (Craver, 2007).

5. BIASES OF ABSTRACT MECHANISM SCHEMAS

The network mechanisms of both sociology and economics are general because they are described in an abstract manner. However, as seen, the sociological models do not aim at unification, in the sense of the most parsimonious set of behavioural assumptions from which the maximum number of consequent phenomena can be deduced. The sociological models aim at a set of mechanism schemas, which are extrapolatable to the extent that the underlying psychological mechanisms are general. Concepts of mechanism are accompanied by heuristics for finding them, and there are grounds for concern that aiming at maximally general, but especially unifying, mechanism schemas at the cost of interest in lower level mechanistic detail may induce biases in social research.

Abstract mechanistic knowledge can be exported to different contexts purely because its application is less demanding: whatever entities satisfy the behavioural assumptions will instantiate the mechanism schema. This exportability however comes with a series of concerns (Kuorikoski, 2009): First, it may be that the 'same' abstract mechanism concept cannot be used in different domains of application with the same level of abstraction. In such cases its use may not allow similar inferences in the different contexts. For example, if we are using the concept of market mechanism to explain how individual interaction leads to macro outcomes in situations in which there is no actual common currency (say, 'the market' for spiritual services in rational choice studies of religion), drawing hasty analogical conclusions dependent on the existence of that currency can be tempting. Second, it is also tempting to infer from the use of a certain mechanism-concept that the constituents of the system so described have other causal features in common with some other, often exemplary, realization of the 'same' mechanism.

Possible examples of the first two sources of error can be found in the possible applications for the *emperor's dilemma model* suggested by Centola et al. They suggest that the emperor's dilemma may be responsible for such things as gossiping about social faux pas by snobs, luxury goods with an anomalous inverted demand elasticity ('luxury fever') and public adoration of a bully by fearful schoolboys. Each suggested phenomenon seems plausibly to be an instance of a publicly enforced but privately disliked norm. But care is to be taken if the emperor's dilemma 'mechanism' is invoked to explain such norms, since luxury fever can also result from simple signalling of high income (Corneo & Jeanne, 1997), and at least some ethnographic studies are hard to square with the hypotheses that signalling of sincerity is the motivating factor in 'dishonest' bullying (Burns et al. 2008). We are not suggesting that these suggested applications of the *emperor's dilemma* mechanism are necessarily erroneous. We only want to draw attention to the apparent ease with which such generalizing claims can be made and the importance of independent empirical evidence for such mechanism attributions.

The third possible bias associated with working with mechanisms as abstract forms of interaction is that the application of a particular mechanism concept to a system may obscure the fact that the system may have other important causal properties besides those that result from that mechanism. This worry is especially pertinent if an abstract mechanism schema is used because of its unificatory promise. The attempt at finding such patterns of action that fit into the unifying schema may crowd out investigation into other, possibly more important, properties and relations between the actors studied. For example, the use of the hypothetical selection of co-authorship relations as an example of the analytical properties of the general modelling framework of Jackson and Wolinsky (1996) may easily lead to the impression that the model is intended to capture 'the economic essence' of individual decisions of entering into such relations. In reality, the strategic weighting of network externalities and effort probably plays only a minor role in the formation of such networks. Sometimes economists treat the fact that a social system can be seen as exhibiting a rational choice network mechanism as an explanatory advance by itself. However, these explanatory triumphs may just be artefacts of the flexibility of the rational choice modelling framework rather than empirical discoveries of new explanatory mechanisms. In such cases, there is no such thing as increase in explanatory understanding.

6. CONCLUDING REMARKS

We have argued that the differences between the strategies of modelling networks in economics and sociology do not derive from different explanatory aims or from a systematic selection of different kinds of properties attributed to the agents. In contrast, we have traced these differences to diverging conceptions of generality, which in turn come with different implications regarding how knowledge from different disciplines is to be integrated.

Field-specific conceptions of epistemic virtues such as generality, which are often only implicit in scientific practice, play a significant role in guiding the way in which models are built, used and evaluated. The influence of such field-specific conceptions may be such that modelling strategies also display significant variations when a common set of modelling tools is utilized for building models intended to serve similar epistemic functions (in the case at hand, for example, for the provision of mechanistic explanations of network phenomena). That a set of tools such as network theory is employed across different sciences does not warrant grand claims of unification (see also Marchionni, 2013). In fact, attention to the use of similar modelling techniques should not lead us to overlook important differences in the way in which such techniques are embedded in disciplinary matrices. Conceptions of epistemic virtues therefore influence transfers of models and modelling techniques across disciplinary boundaries and invites scholars interested in these trends to investigate the conditions under which such transfers promote higher levels of interdisciplinary integration.

Acknowledgments: We thank the editors and two anonymous referees for their helpful comments. Both authors thank the Academy of Finland for financial support.

REFERENCES

Ballester, C., Calvó-Armengol, A. & Zenou, Y. (2006). Who’s Who in Networks. Wanted: The Key Player. *Econometrica 74*, 1403-1417.

Bearman, P. S., Moody, J., & Stovel, K. (2004). Chains of affection: the structure of adolescent romantic and sexual networks. *American Journal of Sociology 110*, 44-91.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: a mechanistic alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences 36*, 421-441.

Burns, S., Maycock, B., Cross, D., & Brown, G. (2008). The Power of Peers: Why Some Students Bully Others to Conform. *Qualitative Health Research* 18: 1704- 1716.

Calvó-Armengol, A., Patacchini, E., & Zenou, Y. (2009). Peer effects and social networks in education. *The Review of Economic Studies 76*, 1239-1267.

Centola, D., Willer, R., & Macy, M. (2005). The emperor’s dilemma: a computational model of self-enforcing norms. *American Journal of Sociology 110*, 1009-40.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology 94*, 95-120.

Corneo, G., & Jeanne, O. (1997). Conspicuous consumption, snobbism and conformism. *Journal of Public Economics 66,* 55-71.

Craver, C. (2007). *Explaining the Brain*. Oxford: Oxford University Press.

Currarini, S., Jackson, M. O., & Pin, P. (2009). An economic model of friendship: homophily, minorities, and segregation. *Econometrica 77*, 1003-1045.

De Martí, J., & Zenou, Y. (2011). Social networks. In I. C. Jarvie, & J. Zamora-Bonilla (Eds.), *The Sage Handbook of The Philosophy of Social Sciences* (pp. 339-361). Sage.

Gilovich, T., Medvec, V. H. & Savitsky, K. (1998). The Illusion of Transparency: Biased Assessments of Others' Ability to Read One's Emotional States. *Journal of Personality and Social Psychology* 75, 332-346.

Goyal, S. (2007). *Connections: An Introduction to the Economics of Networks*. Princeton: Princeton University Press.

Granovetter, M. S. (1983). The strength of weak ties. *American Journal of Sociology 78*, 1360-1380.

Grüne-Yanoff, T., & Schweinzer, P. (2008). The Role of Stories in Applying Game Theory, *Journal of Economic Methodology 15*, 13-46.

Hausman, D. (1992). *The Inexact and Separate Science of Economics*. Cambridge and New York: Cambridge University Press.

Hedström, P., & Ylikoski, P. (2010). Causal mechanisms in the social sciences. *Annual Review of Sociology 36*, 49-67.

Jackson, M. O. (2008). *Social and Economic Networks*. Princeton: Princeton University Press.

Jackson, M. O., & Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of Economic Theory 71*, 44-74.

Kincaid, H. (2012). Some Issues Concerning the Nature of Economic Explanation. In U. Mäki (Ed.), *Philosophy of Economic*s (series *Handbook of the Philosophy of Science,* volume 13) (pp. 137-158). Elsevier.

Kuorikoski, J. (2007). Explaining with equilibria. In Persson, J., & P. Ylikoski (Eds.) *Rethinking Explanation*. Springer.

Kuorikoski, J. (2009). Two concepts of mechanisms. *International Studies in Philosophy of Science 23*, 143-160.

Lawson, T. (2003). *Reorienting Economics.* London and New York: Routledge.

Lehtinen, A., & Kuorikoski, J. (2007). Computing the Perfect Model: Why Do Economists Shun Simulation? *Philosophy of Science 74*, 304-329.

Marchionni, C. (2013). Playing with networks: how economists explain. *European Journal for Philosophy of Science 3,* 331-352.

Marchionni, C., & Ylikoski, P. (2013). Generative explanation and individualism in agent-based simulation. *Philosophy of the Social Sciences 43*, 323-340.

Matthewson, J., & Weisberg, M. (2009). The Structure of Tradeoffs in Model Building. *Synthese 170*, 169-190.

Morgan, M. (2001). Models, stories and the economic world. *Journal of Economic Methodology 8*, 361-384.

Ross, D. (2005). *Economic Theory and Cognitive Science: Microexplanation*. Cambridge MA: The MIT Press.

Savitsky, K., Epley, N., & Gilovich, T. (2001). Do others judge us as harshly as we think? Overestimating the impact of our failures, shortcomings, and mishaps. *Journal of Personality and Social Psychology* 81, 44–56.

Willer, R., Kuwabara, K., & Macy, M. W. (2009). The False Enforcement of Unpopular Norms. *American Journal of Sociology 115*, 451-490.